

Napoléon et l'évaluation bibliométrique de la recherche. Considérations sur la réforme de l'université et sur l'action de l'agence national d'évaluation en Italie

This is the peer reviewed version of the following article:

Original:

Baccini, A. (2016). Napoléon et l'évaluation bibliométrique de la recherche. Considérations sur la réforme de l'université et sur l'action de l'agence national d'évaluation en Italie. CANADIAN JOURNAL OF INFORMATION AND LIBRARY SCIENCE-REVUE CANADIENNE DES SCIENCES DE L'INFORMATION ET DE BIBLIOTHECONOMIE, 40(1), 37-57.

Availability:

This version is available <http://hdl.handle.net/11365/983569> since 2018-09-21T08:44:05Z

Terms of use:

Open Access

The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. Works made available under a Creative Commons license can be used according to the terms and conditions of said license.

For all terms of use and more information see the publisher's website.

(Article begins on next page)

Napoléon et l'évaluation bibliométrique de la recherche. Considérations sur la réforme de l'université et sur l'action de l'agence national d'évaluation en Italie.¹

par Alberto Baccini (Dipartimento di economia politica e statistica, Università di Siena, Italie)

À partir de l'an 2000 le système universitaire italien a fait l'objet de plusieurs *réformes* qui doivent être contextualisées dans le cadre des processus de Lisbonne et de Bologne : tout d'abord la réforme Berlinguer (L. 30 10 février 2000) puis la réforme Moratti (L. 53 28 mars 2003) et enfin la réforme Gelmini (L. 240 du 30 décembre 2010). Ces réformes peuvent être considérées comme plusieurs tentatives pour passer d'un système universitaire contrôlé par le centre de style napoléonien (Ivar Bleiklie et Svein Michelsen, 2013), à un système plus décentralisé, basé sur l'autonomie, l'évaluation de la recherche et de la qualité de l'enseignement.

Les réformes se sont passées dans un climat de plus en plus caractérisé par la prévalence d'une rhétorique de l'urgence pour stopper le déclin de l'université et de la recherche italienne (Moretti 2010). Une rhétorique similaire a été décrite par (Aust and Crespy 2009) relative à la nécessité de rendre plus performatif et plus visible au niveau international le système de la recherche français. L'effondrement de l'université italienne est narrée dans des dizaines d'articles publiés par le *Corriere della Sera*, dans des ouvrages d'investigation et des sites web, comme pour exemple *La voce*, qui est un site géré par des économistes académiques, et par un flot d'articles presque-scientifique (pour une synthèse G. De Nicolao, 2012). Le texte plus significatif de ce déferlement est un pamphlet intitulé *L'Università truquée* (Perotti 2008).² Sur la base de ces sources, « ce que tout-le-monde connaît » sur l'Université et la recherche en Italie est que l'université italienne est « malade » (Regini 2009) ; qu'il s'agit d'un monde dominé par une caste de *barons* (Carlucci and Castaldo 2009) complètement autoréférentiel, votée pour défendre les intérêts de la corporation et de la famille. En plus,

« la partie saine, bien sûr, du corps enseignant non seulement tolère maintenant un comportement qui a toujours été répandu, alors qu'autrefois jugé au moins contestable, mais choisit ses dirigeants parmi ceux qui ont poursuivi ces comportements avec plus d'énergie, d'efficacité et de succès. » (Graziosi, 2010: 15)

Le système de recrutement est corrompu par le familisme et les clientèles locales (Perotti, 2008: 52-66). L'université italienne produit peu et sa recherche est de mauvaise qualité : « Le rôle des universités dans la recherche mondiale est donc marginal » (Graziosi, 2010: 8). Elle forme des diplômés dont les entreprises n'ont pas besoin (Giavazzi 2012), et surtout « l'Italie a trop de diplômés » (Rocca 2014). Le système n'est pas réformable et l'Université est responsable pour sa part du déclin de l'Italie (Gagliarducci et al. 2005). Les ressources dédiées par l'État à l'Université sont utilisées de façon inefficace et sont donc excessives (Perotti 2008). Ce n'est qu'à partir de 2010 que cette vision a commencée à être remise en question par un livre *Les*

¹ Une précédente version de cet article a été présentée à la première journée d'étude « Bibliométrie et indicateurs de la recherche: enjeux et perspectives » organisé à l'université de Toulouse le 11 juin 2013. Je voudrais remercier Guillaume Cabanac et Gilles Hubert pour les commentaires ponctuels et la collaboration pendant l'écriture de ce texte ; je voudrais aussi remercier les deux évaluateurs anonymes pour leurs suggestions.

Déclaration de conflit d'intérêt : L'auteur est membre de la rédaction du blog www.roars.it.

² Pour une critique du pamphlet (Moretti 2009).

chercheurs ne grandissent pas sur les arbres (Sylos Labini and Zapperi 2010), qui fustige le manque de ressources qui caractérise la recherche italienne. Il est suivi par *L'Université maudite* (Coniglione 2011). En octobre 2011 est né le blog collective www.roars.it, qui se propose de « démystifier les mensonges » et de « contribuer à une discussion pensif et compétent » sur l'université et la recherche en Italie.³ Au début de 2015, après moins de quatre ans d'activité, ce blog a eu plus de 10 millions de visites, a publié environ 2 milles articles signé pas plus que 200 auteurs (Redazione Roars 2015). De son fondation, le blog est devenu le point de référence principale pour la discussion sur la politique de l'université et de la recherche en Italie (Reale 2012).⁴

Pendant ce temps, néanmoins, des solutions avaient déjà été mises en œuvre. Une première réponse a été constituée par la création de nouveaux centres de « recherche d'excellence » financés de façon appropriée (IIT institut italien de la technologie, de Gênes ; SUM (haute école en sciences sociales); IMT). Une deuxième ligne de réponse, beaucoup plus complexe, inclut la réforme de la gouvernance des universités et de leur structure organisationnelle. Dans cette réforme, le rôle de premier plan est la création d'une agence centrale qui ambitionne d'évaluer la recherche et la didactique à l'aide d'*indicateurs objectifs*. L'idée se répand que le salut de l'université italienne passe par l'évaluation pour tous (Graziosi 2010), « l'évaluation est la panacée pour tous les maux de l'université » (Capano 2013). Les ressources financières doivent être réparties par le gouvernement en proportion aux résultats de l'évaluation (Graziosi 2010) : « cette allocation doit être impitoyable » (Perotti 2008). « Récompenser le mérite » et « punir le démerite » c'est le nouveau slogan (Perotti 2008). Cette vision a été d'abord adoptée par le centre-gauche. Le ministre Fabio Mussi – ministre du deuxième gouvernement Prodi de mai 2006 à mai 2008 - est l'auteur du premier projet de l'Agence nationale d'évaluation de l'Université et de la recherche (ANVUR). En 2007 il présentait son projet en soutenant : « aujourd'hui la pourcentage des financement orientée de l'évaluation méritocratique est presque nul. L'agence d'évaluation naît pour prendre une photo réelle, pour faire un ranking absolu des universités et pour récompenser les améliorations » (Bonazzi 2007). Ce projet d'Agence a été réalisé sans modifications substantielles par Maria Stella Gelmini, ministre du gouvernement Berlusconi (de mai 2008 à novembre 2011). L'ANVUR en effet est devenu un des centres principaux de la « loi Gelmini » (Rubele 2012). L'évaluation est considérée comme la solution pour trois raisons. La première raison est que tous les pays de l'OCDE adoptent des systèmes plus ou moins complexes d'évaluation de la recherche et d'assurance de qualité de l'enseignement (Rebora 2013). L'évaluation est considérée un système pour vérifier si les ressources sont utilisées efficacement, et pour rendre compte aux contribuables comment leur argent est dépensé (Baccini 2010). La deuxième raison est que l'évaluation permet ainsi d'introduire la compétition pour la répartition des ressources (Capano 2013), car la loi (Décret 180/2008, art. 21) prévoit explicitement qu'une pourcentage croissant du financement publique soit répartie sur la base des résultats de l'évaluation (pour une reconstruction analytique de ce point, on peut voir Banfi and Viesti 2015). La troisième, enfin est qu'elle permet de distinguer *objectivement*, grâce à la bibliométrie, la bonne recherche et la recherche autoréférentielle. En particulier, la bibliométrie est considérée comme un outil pour mettre en crise le clientélisme des compétitions pour le concours de titulaire à l'université, que est généralement considéré

³ <http://www.roars.it/online/chi-siamo/organizzazione-e-scopo-del-sito/>

⁴ L'un des membres du Conseil de l'Agence National d'Evaluation a soutenu que le modèle du site www.roars.it est celui de *l'advocacy* : « fondé sur des commentaires rapides, des jugements sommaires et des attaques personnelles. [ROARS a] besoin de convaincre les institutions et les parties prenantes de ses motifs. Ils recherchent des contacts avec toutes les parties, tentent d'influencer les programmes pour les élections de 2013, ils cherchent la visibilité sur les médias et dans les institutions de la science » (Bonaccorsi 2013). Probablement, l'expérience du blog peut être considérée plus précisément comme un cas de naissance d'un « *scientific counterpublic* » (Hess 2012).

depuis longtemps (Garbisa and Calzà 1995; Calza and Garbisa 1995; Checchi 1999) le problème principale de l'université italienne.

Dans les pages qui suivent, après une description de la structure organisationnelle de l'université italienne, nous présentons la structure institutionnelle de l'Agence d'évaluation et les modalités de mise en œuvre de ses premières activités. Sur cette reconstruction nous allons développer des observations critiques soit sur la conception de la gouvernance de l'ANVUR, soit sur la réalisation des activités d'évaluation.

1. Une structure institutionnelle déséquilibrée

La structure de l'université italienne est qualifiée de napoléonienne (Bleiklie et Michelsen, 2013) parce qu'elle est caractérisée par le contrôle centralisé et bureaucratique exercé par le ministère. L'académie italienne a discuté longuement de l'autonomie des universités, prévue à l'article 33 de la Constitution (Barbati 2013). Les universités étaient gouvernées par les Sénats académiques et les conseils d'administration, élus par la communauté universitaire. Le système a été organisé depuis 1990 autour des facultés gérant les enseignements et les départements qui organisent la recherche. Le financement de l'Université était centralisé et reposait essentiellement sur des indicateurs dimensionnels et très peu sur les résultats en recherche et en enseignement. Le financement des projets de recherche a toujours été octroyé en ruisselets non coordonnés. La représentation au niveau central de la communauté universitaire est attribuée au Conseil Universitaire National (CUN), élu par la communauté universitaire (Loi 16 Janvier 2006). Les Recteurs des universités sont organisés au sein d'une conférence nationale (CRUI) (Capano 2013).

La loi Gelmini en 2010 a profondément modifié cette structure de gouvernance de l'Université, poussant vers une gestion des universités centrée sur le Recteur (élu) et le Conseil d'administration (nommé par le recteur et composé aussi de personnels non-académiques), avec un rôle très mineur des sénats universitaires. Cette autonomie accrue théorique nécessite au moins une évaluation centralisée du travail local (Capano, 2010). Le ministère a créé pour cette tâche l'ANVUR, qui opère depuis le 22 février 2011, dessinée sur le modèle de l'AERES française. À l'ANVUR sont attribuées les fonctions d'évaluation de la recherche et celles de « l'assurance de qualité » de l'enseignement supérieur, et plus récemment aussi d'évaluation de la structure administrative des universités (decret n.90/2014). Le Conseil de direction de l'ANVUR est composé de sept membres avec des responsabilités opérationnelles directes, qui sont nommé directement par le ministre. On peut considérer tout ça comme une forme de re-centralisation du système académique italien, qui transforme la centralisation précédente à travers une centralisation manageriale exercée à travers l'ANVUR. Une dynamique similaire de centralisation du système politico-administratif britannique a été documentée par Hoggett (1996). Bezes (2005) a soutenu que l'application de réformes managériales en France a abouti à conforter le pouvoir des hauts fonctionnaires qui maîtrisent notamment la définition du processus évaluatif. En plus, la loi Gelmini n'a pas modifié la structure du CUN, mais a posé les bases de son un progressive « affaiblissement fonctionnel » (Sandulli 2011). En particulier sa structure et sa composition⁵ sont considérée désormais obsolète, et ses procédures décisionnelles trop lentes en référence à la nouvelle gouvernance des universités (Barbati and Lenzi 2011; Lenzi and Barbati 2011). En plus la loi Gelmini a soustrait au CUN la compétence sur les questions disciplinaires des professeurs, en l'attribuant directement aux universités (Pezzella 2011). Aussi le système des financements des projets de recherche n'a pas été modifié, en

⁵ Le CUN est composé de 58 membres qui ne représente pas seulement les professeurs et les chercheurs des universités (42 membres), mais aussi les étudiantes (8 membres), le personnel administratives des universités (3 membres), la Conference des Recteurs (3 membres), le faculté universitaires (1), les directeurs généraux des universités (1). Les 42 représentantes des professeurs sont élus en nombre de trois pour chacune des 14 Area de recherche; pour chaque Area doivent être élus un *professore ordinario* ; un *professore associato* et un chercheur.

dépendant principalement directement du Ministère de l'Université, et pour lignes spécifiques de recherche, d'autres ministères (Santé, Agriculture etc.). Le résultat de cette réforme est un système déséquilibré (Travi 2011), c'est-à-dire un système qui n'a pas de contrepoids (*checks and balances*) qui serve au contrôle mutuel des acteurs principaux (Baccini 2013; Ammannati 2013). En particulier, il convient de souligner l'absence d'un organisme centralisé pour le financement de la recherche et d'un organisme clairement délimité représentatif de la communauté académique. Le système actuel finit par attribuer un rôle centrale à l'ANVUR non seulement sur les procédures d'évaluation, mais aussi sur toutes les procédures pour l'assignement des financements aux universités (Travi 2011). Capano a défini l'ANVUR une "organization monstre dont les compétences sont sans précédent dans les agences du monde occidental" (Capano 2013). Les conséquences de ces choix institutionnelles sont analysées selon différentes perspectives. Dans une première perspective, l'ANVUR est considérée comme le centre d'un système où le Ministre n'a plus le contrôle de la politique de la recherche et de l'instruction universitaire, parce que il se limite à ratifier les décisions prises par l'organisme technique d'évaluation (Capano 2013). Une seconde perspective considère plutôt comme un problème réel (Baccini 2013) ou seulement potentiel (Della Cananea 2011) la possibilité de la politique d'exercer un contrôle direct sur la recherche et l'université en nommant les membres du conseil d'administration de l'ANVUR. Une perspective plus radicale considère l'évaluation et l'existence d'une agence nommée par le ministre en conflit avec la liberté de recherche et d'enseignement définies dans la Constitution italienne (Pinelli 2011), ou comme un outil d'espionnage pour le contrôle de l'université (Pinto 2013). Plus faibles critiques proviennent aussi de celui qui ont défini la structure institutionnelles de l'ANVUR, et qui considèrent les problèmes de l'agence comme causée par une mauvaise choix des composants du conseil de direction de part du ministre (Modica 2012). Dans ce débat, la défense du *status quo* a été proposée par un membre du conseil d'administration de l'ANVUR, qui a soutenu l'indépendance complète de l'agence de la politique (Bonaccorsi 2013).

Les conséquences de cette structure déséquilibrée sur la mise en œuvre des pratiques d'évaluation peuvent être illustrées en décrivant comment le Ministère et l'ANVUR ont conçu et réalisé l'évaluation nationale de la recherche, et quelques aspects de l'habilitation scientifique nationale.

2. L'exercice nationale d'évaluation de la recherche (VQR 2004-2010)

En 2003, le *Comitato di Indirizzo per la Valutazione della Ricerca* (CIVR: Comité d'Orientation pour l'Évaluation de la Recherche) avait mené le premier exercice d'évaluation de la recherche en Italie (VTR) après 15 ans de la première expérience impulsée par Margaret Thatcher au Royaume-Uni en 1986 (RAE). L'objectif était d'évaluer l'*excellence* des activités de recherche menées par les universités et les instituts de recherche publics supervisés par le MIUR. Le VTR était conçu sur l'exemple du RAE britannique. Le VTR concernait 20 domaines de recherche au cours de la période 2001-2003 et 102 institutions de recherche dont 77 universités, 12 organismes publics de recherche et 13 organismes privés de recherche. Les productions sélectionnées étaient au nombre de 18 508 ; 17 309 étaient évalués par 6 661 experts,⁶ dont 1 465 étrangers ou résidents à l'étranger avec un « coût direct du processus » de 3,55 millions d'euros.⁷ En un mot: ce fut une expérience avec un impact modeste, coûteuse, avec quelques défauts évidents qui devront être surmontés dans les exercices suivants (Reale 2008; Vaira 2011). À partir de 2010 les résultats de ce premier exercice d'évaluation sont utilisés pour répartir un pourcentage (7%) du financement public aux universités.

⁶ http://vtr2006.cineca.it/php5/relazione_civr/output/totale.pdf

⁷ <http://www.miur.it/Seminario/PDF/Cuccurullo.pdf>; <http://www.analysis-online.net/wp-content/uploads/2013/03/cuccurullo.pdf>

Seulement, en juillet 2011, cette expérience a été suivie par la *Valutazione della Qualità della Ricerca* (VQR : Évaluation de la Qualité de la Recherche) 2004-2010 dirigé par l'ANVUR. La VQR s'inscrit dans un contexte radicalement modifié, puisqu'elle se déroule avec d'autres procédures qui ont l'ANVUR pour protagoniste: évaluation de la didactique universitaire (AVA), accréditation des doctorats et habilitation scientifique nationale (ASN). La VQR a eu un long délai de gestation, dont le résultat final a été le dessin d'une procédure où les responsabilités du ministre et de l'ANVUR apparaissent en grande partie superposées, probablement à cause des interactions, qui ne sont pas documentées par des actes publics, entre le ministre et l'ANVUR pendant la préparation du décret. Pour illustrer cette superposition des rôles technique et politique, on peut souligner que le décret ministériel a dicté directement les méthodologies et les outils d'évaluation (DM MIUR n. 17 du 15 Juillet 2011), tandis que l'ANVUR s'est limitée à faire des ajustements⁸ à l'orientation ministérielle. Et on peut souligner aussi que c'était l'un des membres du Conseil d'administration de l'ANVUR à déclarer à un quotidien national les objectifs politiques de l'évaluation : « toutes les universités devront recommencer à zéro. Et une fois l'évaluation terminée, nous aurons la distinction entre *researching* [sic] *university* et *teaching university*. À certains nous dirons : vous faites juste le cycle Bachelor, et quelques universités seront fermées » (Fiori 2012). Cette inversion des rôles a été l'objet d'une interrogation à la Ministre par des membres de la Chambre des Députés qui demandait de « définir avec précision le rôle et le domaine d'intérêt » de l'ANVUR.⁹

Pour mieux comprendre la VQR, il est nécessaire de la présenter sommairement.¹⁰ La VQR est caractérisée par un dessin méthodologique hybride qui rend impossible la classer selon les typologies proposées par Hicks (2010). Comme dans le RAE/REF Britannique, il a été décidé d'évaluer la meilleure production de chaque auteur : trois des produits de recherche pour chaque chercheur, ou six pour les chercheurs qui ne sont pas chargés de la didactique. Contrairement au RAE/REF où seulement un petit sous-ensemble de chercheurs participe à l'évaluation, dans la VQR tous les chercheurs sont tenus de participer. La VQR a donc mis en œuvre le choix le plus raffiné : l'évaluation ponctuelle de la qualité de chaque produit de recherche. Différemment du RAE/REF les produits de la recherche sont évalués avec l'un des deux méthodes définies par le décret ministériel : l'évaluation par les pairs, ou l'analyse des citations. Les résultats finaux sont présentés pour les universités et pour les départements, considérés comme la somme des produits présentés par les chercheurs qui y travaillent. Les données individuelles (le score de chaque produit) ne sont pas publiées, mais communiquées en privé à chaque chercheur.

La VQR a évalué 184 742 produits de la recherche, partagés en 14 sujets, comme indiqué dans le Tableau 1, qui correspondent aux 14 *Area* autour desquelles est organisée historiquement l'université italienne.¹¹ Les *Area* 1-9 et 13 sont définies par ANVUR « bibliométriques » parce que l'évaluation est conduite en prévalence avec des indicateurs bibliométriques, et les autres comme « non-bibliométriques », parce que l'évaluation est conduite exclusivement par les pairs.

Pour chaque sujet le décret ministériel a prévu qu'il y aurait un groupe d'experts nommé par l'ANVUR (GEV : groupe d'experts d'évaluation) chargé de la responsabilité de l'évaluation, et que chacun de ces groupes aurait un président choisi parmi les membres. L'ANVUR a préféré nommer en premier les présidents de chaque

⁸ ANVUR, Bando di partecipazione, 7 novembre 2011:

http://www.anvur.org/attachments/article/122/bando_vqr_def_07_11.pdf;

⁹ "Interrogazione a risposta in commissione 5/06193 presentata da Ghizzoni Emanuela (partito Democratico)", 16/02/2012. http://dati.camera.it/ocd/aic.rdf/aic5_06193_16;

http://www.senato.it/documenti/repository/commissioni/comm07/documenti_acquisiti/Prof.%20Benedetto%2029.02.12.pdf.

¹⁰ Pour une discussion plus complète en anglais on peut voir (Abramo and D'Angelo 2015; Baccini and De Nicolao 2015).

¹¹ https://www.cun.it/uploads/storico/settori_scientifico_disciplinari_english.pdf

GEV, puis ensuite les membres restants des GEV pour un total de 450 personnes.¹² Ce mécanisme de nominations en cascade a donné lieu à des problèmes d'équité (*fairness*) dans la composition des GEV. Par exemple, dans les disciplines économiques (Area 13), presque tous les membres du GEV sont coauteurs entre eux ; et bien sept des membres du GEV étaient partisans d'un petit parti politique libériste nommé « Fare per fermare il declino » (Baccini 2014).

Chaque produit de recherche a été évalué par un score qui est défini directement dans le décret ministériel. L'idée à la base de cette partie du décret est probablement une adaptation des « quality profiles » adoptés dans le RAE britannique.¹³ Le décret soutient qu'il y a une distribution statistique partagée au niveau international de la qualité des produits de la recherche pour chaque domaine: les produits qui sont dans la moitié inférieure des scores (50 %) sont de qualité limitée (L: 0 point); les 10 % suivants sont de qualité acceptable (A : 0,2 point), les 20 % suivants sont de bonne qualité (B : 0,50 point), et les 20 % restants sont excellents (E : 1 point). Un chercheur qui présente moins de trois produits est pénalisé (-0,5 pour chaque produit manquant). Les cas de plagiat sont pénalisés avec un score négatif (-2). Cette idée de prendre comme référence pour les jugements « l'échelle de valeur partagée par la communauté scientifique internationale » est vide de sens sur le plan opérationnel, car une telle échelle n'existe tout simplement pas (McNay 2011). Pour cette raison, l'ANVUR a modifié le décret ministériel définissant les critères d'identification des niveaux d'excellence des produits. Malheureusement, ces critères ne sont également pas bien formulés et n'ont pas une signification claire d'un point de vue opérationnel.¹⁴ Pour exemple la définition de « produit excellent » est formulée comme une tautologie: « les produits de niveau excellent sont ceux qui sont reconnus internationalement comme excellents pour l'originalité, la rigueur méthodologique et la pertinence interprétative ».¹⁵

Au-delà des problèmes de définition, dans la pratique les scores sont attribués aux produits de la recherche avec un mélange d'instruments qu'il est difficile aussi de synthétiser. Pour les secteurs non-bibliométriques le critère principal est l'utilisation de l'examen des pairs informés (informed peer review : IR), c'est-à-dire les pairs ont à leur disposition les meta-informations sur les produits de recherche évalués, et aussi sur le classement des revues où ils sont publiés. Le domaine de l'économie/gestion/statistiques (Area 13) a constitué une exception parce les articles sur revues ont été évalués exclusivement sur la base du classement bibliométrique des revues où ils sont publiés ; ce classement a été développé directement par les membres du GEV.¹⁶

Pour les secteurs bibliométriques, les critères bibliométriques sont utilisés prioritairement. Mais la bibliométrie varie non seulement entre les différents secteurs, mais même entre sous-ensembles de disciplines à l'intérieur d'un même secteur. Le tableau 2 montre que l'évaluation VQR est organisée en 44 sous-groupes disciplinaires qui utilisent différentes bases de données, qui agrègent 47 listes des revues différentes en utilisant différentes méthodes (bibliométriques et non-bibliométriques).

En substance, on procède selon une procédure bi-étagée basée sur des « matrices de correspondance entre les classes initiales des revues (colonnes) et des citations (lignes) ». Sur l'axe vertical de la matrice il y a le nombre de citations reçues par les articles ; sur l'axe horizontal l'indicateur bibliométrique qui se réfère aux revues, généralement l'impact factor (IF). Chaque article est classifié en considérant le nombre de citations

¹² Le fichier avec les noms des présidents des GEV, publié sur le site de l'agence, est daté du 10 Octobre 2011. Les noms des membres de la GEV ont été rendus publics le 12 Décembre 2011.

¹³ <http://www.rae.ac.uk/results/intro.aspx>

¹⁴ Critiques similaires aux définitions du RAE britannique ont été soulevées par (McNay 2007).

¹⁵ http://www.anvur.org/attachments/article/122/bando_vqr_def_07_11.pdf

¹⁶ Sur Area 13 on peut voir (Baccini and De Nicolao 2015; Bertocchi et al. 2015; Baccini 2014)

qu'il a reçu et l'impact factor de la revue ou il a été publié. La matrice est partagée en 16 rectangles identifiés en considérant les distributions des IF des revues et des citations selon les critères mentionnés ci-dessus (L : 50% ; A : 10% ; B : 20% ; E : 20%). A chaque article est attribué une classification (E, B, A, L) qui dépend de son placement dans un des rectangles de la matrice. Si les deux informations sont cohérentes entre eux, alors la classification est univoque : si un article est parmi les 20% des articles plus cités de son domaine de recherche, et s'il est publié dans une revue avec un IF qui est dans le 20% des revues avec le plus haut IF, alors l'article est évalué comme E(xcellent). Quand les deux informations ne sont pas cohérentes, c'est le GEV à décider les critères d'évaluation. En quelques cas c'est l'information sur la citation qui prévaut : par exemple quand un GEV a décidé que on doit classer comme « B » les articles qui sont dans les 40% des articles plus cités de leur domaine, mais sont publiés dans une revue classifiée comme acceptable. En d'autre cas est le classement de la revue à prévaloir. En d'autre cas, enfin, le GEV les informations bibliométriques sur un article sont considérés par le GEV si contradictoires - typiquement quand un article reçoit beaucoup de citations, mais il est publié sur une revue de « qualité » limitée - qu'il est nécessaire que l'article soit examiné et classé par des pairs informés (*informed peer review*). Les règles de classification des rectangles de la matrice ont été décidées par les GEV et sont différents parmi les GEV, et aussi pour les différents sous-domaines de recherche à l'intérieur du même GEV. L' ANVUR a développé, comme on peut lire dans le Tableau 2, 48 matrices différentes.

L'utilisation de ce mécanisme soulève au moins deux types de problèmes. Le premier est d'ordre pratique : il est difficile pour un chercheur de choisir les produits de recherche à soumettre à l'évaluation. Un groupe de travail de l'Université de Naples a produit un document pour aider les chercheurs des secteurs bibliométriques à sélectionner leurs produits, en limitant l'analyse à la seule base de données Web of Science (WoS). Le document se compose de 189 pages.¹⁷ Le deuxième problème est le plus grave : le décret ministériel a fixé la probabilité d'un article d'être jugé comme E, B, A, L de manière uniforme pour tous les domaines et sous-domaines de recherche; les matrices élaborées par les GEV n'ont pas respectées cette uniformité. La première raison de ça est que les GEV ont adopté des seuils différents pour la distributions des IF des revues. Pour exemple le GEV09 (ingénierie) a utilisé des matrices construites de manière avec une partition en quartiles des IF des revues, en déterminant des scores 1,4 fois plus élevés que dans toutes les autres disciplines (Baccini and De Nicolao 2012). La deuxième raison est que, comme on peut voir dans le tableau 2, les critères et les bases de données utilisées pour classer les revues sont multiples et beaucoup différents parmi les GEV. La troisième raison est d'ordre statistique. On sait à partir au moins de (Seglen 1997) que la corrélation entre le facteur d'impact de la revue et le nombre de citations qu'elle a publiés est faible (Lozano et al. 2012), et surtout elle n'est pas la même en tous les domaines de recherche. Ça signifie que les dimensions des rectangles construits à l'intérieur des 48 matrices ne sont pas homogènes, puisque elles dépendent de la distribution conjointe des citations et de l'IF des revues.¹⁸ En certains domaines a été plus facile d'être jugé excellent simplement parce-que les revues avec un haut IF ont publié plus d'articles dans la période considérée. Pour tout ça, ANVUR a recommandé de ne pas comparer les résultats de la VQR entre GEV différents (ANVUR 2013).¹⁹

On doit noter, enfin, les conséquences de ces différences sur la possibilité d'utiliser les résultats de la VQR pour comparer universités et départements. Les scores finaux calculés par ANVUR pour une université ou un département dépendent de leur composition en termes de domaine scientifique du personnel évalué: un

¹⁷ <http://www.crui.it/HomePage.aspx?ref=2071#>

¹⁸ Le GEV de AREA 2 (physique) a calculé les probabilités des articles d'être évalués comme Excellent etc. en considérant les distributions conjointes des IF et des citations pour les sub-domaines de la physique (ANVUR 2013) .

¹⁹ Dans l' « Appendice A » du *Rapporto finale*, ANVUR a publié aussi la dimension de la distorsion sur les résultats finaux des différents GEV (ANVUR 2013).

département avec plus des chercheurs en un secteur où a été plus facile avoir des scores élevés tends à performer mieux qu'un département avec plus chercheurs en secteurs où le score moyenne a été plus bas. En ce cas, ANVUR n'as pas retenu de recommander des catégories et as publié les *ranking* des universités et des départements comme un petit livre (AA.VV. 2013) distribué avec un quotidien national (*Le Corriere della sera*). La Conférence des Recteurs des universités a opéré une tentative de corriger les distorsion en développant un modèle alternative pour calculer les scores finales des départements,²⁰ qui a été en tous cas critiqué (G. De Nicolao 2014). Malgré ces problèmes, pour l'an 2014, les résultats de la VQR sont utilisés pour distribuer une partie (12,6%) du financement publique à l'université.²¹

Utiliser une grande variété d'indicateurs bibliométriques, avoir laissé le choix des critères bibliométriques pour la détermination des seuils à chaque sous-groupe d'experts, coupler l'utilisation de ces indicateurs à l'examen par les pairs, donne lieu à des résultats qui sont d'une fiabilité douteuse à tous les niveaux d'agrégation.²² La variété des métriques va à l'encontre d'une homogénéité de la méthodologie d'évaluation des produits de la recherche, la seule qui permette l'agrégation non-biaisée des résultats et la comparaison entre les disciplines. De plus, la non publication des résultats individuels ne permette pas le contrôle des résultats par la communauté scientifique qui ne peut pas reproduire les calculs réalisés par l'Agence d'évaluation pour grouper les résultats individuels.

3. L'habilitation scientifique nationale : la bibliometrie et les procedure administratives

La loi 240/2010 art. 16 a établi l'Habilitation scientifique nationale qui sert à définir la qualification scientifique comme condition nécessaire pour accéder au grade de professeur, avec ses deux échelons : *ordinario* et *associato*.²³ La procédure est régie par le DM 76/2012 qui a introduit l'idée que les membres du jury de recrutement et les candidats doivent « dépasser les médianes ». La médiane divise en deux parties égales la population des candidats ou des professeurs qui sont éligibles comme membres du jury de recrutement. Une fois de plus, la distinction entre secteurs bibliométriques et non bibliométriques a été introduite dans la procédure. Pour les secteurs bibliométriques, obtenir l'habilitation nécessite de dépasser deux médianes sur trois, calculées sur le nombre de produits scientifiques, de citations reçues, et du *contemporary h-index*, c'est-à-dire une des variantes du *h-index* (Egghe 2010). Pour les secteurs non-bibliométriques, il faut dépasser une médiane sur trois calculées sur le nombre de monographies, le nombre d'articles et de chapitres de livres, et le nombre de publications dans des revues classées excellentes – classe A. Pour réaliser tout ça, ANVUR a dû développer des listes des revues scientifiques, et des liste des revues de Classe A.

Avec ce système la bibliométrie a été introduite massivement dans les procédures administratives pour l'habilitation. Cette choix a été saluée par des commentateurs comme un tournant pour l'université italienne (parmi les autres Massarenti 2012); membres et collaborateurs de l'ANVUR ont justifié le mécanisme des « médianes » comme une « explosion de méritocratie » parce-que il tends « vers l'amélioration lent et progressive de la qualité scientifique des professeurs » (Graziosi 2012; Benedetto 2011). Cette choix a aussi soulevé beaucoup des critiques (pour exemple Banfi 2013); le juge constitutionnel Sabino Cassese a synthétisé ces critiques en soulignant que l'usage de la bibliométrie à l'intérieur des processus

²⁰ http://www.anvur.org/index.php?option=com_content&view=article&id=609&Itemid=563&lang=it

²¹ <http://attiministeriali.miur.it/anno-2014/novembre/dm-04112014.aspx>; en 2013 la pourcentage des ressources distribuées sur la base des résultats de l'évaluation a été un 8% <http://attiministeriali.miur.it/anno-2013/agosto/dm-08082013.aspx>.

²² Ce problème est amplifié par l'utilisation de la bibliométrie et de la IR. Sur ce point on peut voir (Baccini and De Nicolao 2015).

²³ Pour une encadrement des problèmes du recrutement des professeurs en Italie (Banfi 2015).

administratifs en se prêtant à des recours et contentieux judiciaires, détermine le passage de la responsabilité finale sur les procédures d'évaluation de la communauté scientifique aux tribunaux administratifs (Cassese 2013). Cette critique, formulée au début des procédures, a été vérifiée par les événements successives : un contentieux énorme s'est développé autour de l'habilitation à propos des classement des revues, et surtout à propos des décisions prises par les commissions d'habilitation. Les tribunaux administrative ont produit plus que 2 500 document en concernant l'habilitation scientifique national, dont environ 900 sentences.²⁴ L'épisode le plus important a été l'annulation de tous les résultats de l'habilitation pour le secteur du droit privé.²⁵

Dans le cadre de l'habilitation, l'action de l'ANVUR a rencontré des problèmes concernant le calcul des médianes, l'utilisation des bases de données bibliographiques et la classification des revues. En particulier en ce qui concerne les médianes, l'ANVUR a publié les valeurs des médianes et les règle de calcul le 13 août 2012. Après deux semaines, précisément le 27 août 2012, l'ANVUR a retiré ces médianes, et a publié les médianes corrigées avec une nouvelle règle de calcul. Ces nouvelles médianes ont déplacé vers le haut tous les secteurs bibliométriques et vers le bas tous les secteurs non-bibliométriques. Le site www.roars.it, après avoir tenté sans résultat d'inférer les raison plausible de ce changement (G. De Nicolao 2012), a soulevé publiquement le problème en publiant dix demandes pour l'ANVUR (Redazione Roars 2012b). Le Conseil Universitaire National a ensuite également demandé des explication directement au Ministre de l'Université (Lettre CUN n. 1474 du 12/9/2012).²⁶ Les réponses de l'ANVUR ont enfin expliqué les erreurs et les incertitudes, en attribuant la responsabilité au ministre et à l'insuffisance des bases de données utilisées.²⁷

Comme souligné auparavant, l'ANVUR a mis en place rapidement (1) des listes de revues scientifiques pour l'Habilitation et (2) des listes de meilleures revues pour chaque secteur non-bibliométrique (classe A). ANVUR a adopté une définition de revue scientifique sans considérer la discussion (Pontille and Torny 2010) et les règles conventionnelles adoptées au niveau international, comme par exemple la définition contenue dans le *Manuel de Frascati* (OECD 2002). En particulier ANVUR a défini comme « revue scientifique » une revue qui ait « un ou plusieurs des éléments suivants :

- une description de la revue ou de sa politique rédactionnelle avec une référence explicite au caractère scientifique de la revue publiant des résultats originaux ;
- un Comité scientifique ;
- un Comité de rédaction dont la composante académique est pertinente et/ou le directeur de la revue a une affiliation universitaire ;
- un processus d'examen des manuscrits ;
- un référencement dans les principaux répertoires nationaux et internationaux de revues de sciences humaines et sociaux ;
- un index et/ou des articles confirmant que ce sont [...] des contributions scientifiques ».²⁸

²⁴ Données élaborées le 26/06/2015 sur le site <https://www.giustizia-amministrativa.it>

²⁵ "Decreto direttoriale recante l'annullamento dei lavori della tornata 2013 della commissione per il settore concorsuale 12/aA1-Diritto Privato", 16/12/2014.

²⁶ http://www.teoretica.it/wp-content/uploads/2012/09/doc_2012-09-12.pdf

²⁷ http://www.anvur.org/attachments/article/253/mediane_spiegate_definitivo_14_settembre_2012.pdf

²⁸ ANVUR, "Documento di accompagnamento: mediane dei settori non bibliometrici".
http://www.anvur.org/attachments/article/253/Documento%20di%20accompagnamento_mediane%20settori%20non%20bibliometrici.pdf

ANVUR a aussi défini les types des revues qui ne devait pas être considérées comme scientifiques.²⁹ L'ANVUR a donc publié des listes de revues « scientifique » dont une grande partie diverge de la définition ci-dessus. Le problème a été soulevé par le site www.roars.it qui a dédié une trilogie d'articles pour documenter les incongruences des listes obtenues en documentant que toutes sortes de revues que l'ANVUR prévoyait expressément d'exclure des listes y sont bel et bien présentes dans la liste finale. ANVUR a en effet classifié comme « scientifique » des quotidiens, des hebdomadaires, magazines spécialisés destinés aux professionnels des industries et des services, revues de divulgation scientifique, revues publiés par des partis politiques (Redazione Roars 2012a; Banfi and De Nicolao 2012; Baccini 2012). Ces listes ont été ensuite stigmatisés par des articles publiés sur les journaux nationaux (Zunino 2012; Amabile 2012; Stella 2012) et aussi sur la presse internationale (Mazzotti 2012). Le président de l'ANVUR a soutenu que les erreurs de classification ont été le résultat d'une procédure qui devait être conclue rapidement, et de la mauvaise habitude des professeurs des universités d'inclure dans les archives des documents des universités tous leurs publications, pas seulement les publication scientifiques (Fantoni 2012).³⁰ Dans un document officiel ANVUR a défendu ses listes comme le résultat d'un « méritoire travail d'élagage ».³¹ De toute façon, ensuite à cette polémique, ANVUR a partiellement modifié les listes,³² mais ANVUR a aussi dû modifier des classement des revues en force des décisions des tribunaux administrative, activé par les directeur des revues.³³ Selon Massimo Mazzotti :

“The “crazy lists” affair is only a symptom of a deeper malaise. For one thing, the design of Anvur’s semi-mechanized system for the management of careers across public universities reveals a dramatic lack of confidence in the academic sector as a functioning entity. ... The “crazy lists” affair cannot be easily written off as the result of carelessness or malpractice. We should see it as a cautionary tale. It illustrates what can happen when responsibility and autonomy are squeezed out of academic life, and expert decision-making in matters of recruitment, promotion and funding is replaced by a system of explicit and apparently obvious rules.” (Mazzotti 2012)

CONCLUSION

Les lignes de la réforme universitaire entreprise en Italie ces dix dernières années, en particulier la loi 240/2010, se sont inspirées de l'illusion de l'existence des *first best institutions* (Rodrik 2008) : des institutions efficaces quelques soient les conditions et contextes. La révision profonde de la gouvernance de l'Université a été motivée par l'idée qu'introduire de la concurrence entre les universités conduirait à des gains de productivité et à l'améliorer de la recherche et de la didactique (Capano 2013). L'autonomie des universités a été considérée comme la première étape nécessaire pour la concurrence réalisée à travers

²⁹ “Documento di accompagnamento: mediane dei settori non bibliometrici”

http://www.anvur.org/attachments/article/253/Documento%20di%20accompagnamento_mediane%20settori%20non%20bibliometrici.pdf

³⁰ La même justification a été répétée par un des responsables des processus de classification des revues pour les sciences politique (Ferrera 2013) ; un autre collaborateur de l'ANVUR a soutenu que « le scandale » n'est pas que ANVUR ait classifié erronément les revues, mais que « des professeurs écrivent sur ces revues » (Checchi 2013).

³¹ « La classificazione delle riviste nell'ambito dell'abilitazione nazionale », 5 octobre 2012

http://www.roars.it/online/wp-content/uploads/2012/10/chiarimenti_riviste_scientifiche.pdf

³² Une nouvelle version des listes des revues a été publiée le 1 mars 2013 sur le website de l'ANVUR; après quelques heures les listes ont été retirées « à cause d'un erreur » et ont été publiées de nouveau le 6 mars 2013. Aussi ces épisodes ont été documentés sur le site www.roars.it (Giuseppe De Nicolao 2013; G. De Nicolao et al. 2013).

³³ Pour quelques exemples on peut voir ici:

http://www.anvur.org/index.php?option=com_content&view=article&id=254&Itemid=315&lang=it.

l'adoption de mécanismes de financement qui récompensent la qualité de la recherche et de la didactique. L'évaluation a été considérée dans ce contexte comme l'élément clé du système, car il peut simuler le fonctionnement des marchés (Agasisti and Catalano 2006). L'évaluation par les pairs, jugée arbitraire, est alors remplacée par des indicateurs « objectifs » qui doivent guider la distribution des ressources selon la règle : « plus des ressources a les universités meilleurs » (Perotti 2008).

L'introduction de ces innovations dans un appareil administratif napoléonien, a généré un système institutionnel complètement déséquilibré. Le système napoléonien avait finalement octroyé aux professeurs un degré élevé d'indépendance tout en préservant l'autonomie de la recherche (Bleiklie and Michelsen 2013). L'utilisation opportuniste de cette autonomie a encouragé l'hostilité publique mentionnée au début de cet article, qui a été le prétexte à l'action législative de la dernière décennie. Le centre de gravité du système est passé de la ministre et de l'administration ministérielle à l'Agence nationale d'évaluation de l'Université et de la recherche. Une agence qui exerce toutes les activités d'évaluation de la recherche, de l'appareil administratif des universités et l'assurance de la qualité de l'enseignement. Le raccordement de ces nouvelles structures sur un corps napoléonien a fini par renforcer le rôle du Centre en attribuant directement à la politique et à le groupe d'élites autour l'ANVUR choisis par le gouvernement, un pouvoir sans contrepoids sur la recherche et les universités. Une des conséquences de cette nouvelle architecture est probablement le changement de la géographie du pouvoir dans les universités :

“The new rules are able to modify the composition of the actors influencing the knowledge production in universities, with the reinforcement of those individuals more involved in peer reviewing and bibliometric activities, and groups more capable to colonize the disciplinary peer review committees. It is uncertain the extent to which the mentioned transformation of individual professional habits would produce the effect of reinforcing the universities [...] or just re-adjust the internal power distribution” (Reale 2012).

L'État italien dépense désormais des ressources considérable pour l'évaluation de la recherche et de l'enseignement supérieur (Baccini et al. 2013). Un des slogans adopté pour justifier les ressources dédiées pour l'évaluation est que « une évaluation imparfaite c'est mieux que l'absence d'évaluation ». ³⁴ Les erreurs et les discussions qui ont accompagné les activités de l'ANVUR ont probablement miné en Italie le degré de légitimité des processus d'évaluation. On peut considérer la récente expérience italienne comme un cas exemplaire « d'utilisation anarchique pour ne pas dire sauvage » de la bibliométrie, selon l'expression de (Gingras 2008). Sans corrections les ressources dépensées pour l'évaluation n'apparaissent pas justifiées en termes d'améliorations futures. Il n'est pas démontré que l'évaluation améliore la recherche. Une mauvaise évaluation, telle que celle mise en œuvre par l'ANVUR, peut avoir des effets négatifs sur la qualité de la recherche, car elle risque de modifier les comportements des chercheurs de manière imprévisible et indésirable.

Bibliographie

- AA.VV. (2013). *I voti all'università. La Valutazione della qualità della ricerca in Italia*. Milano: Corriere della Sera.
- Abramo, G., & D'Angelo, C. A. (2015). The VQR, Italy's second national research assessment: Methodological failures and ranking distortions. *Journal of the Association for Information Science and Technology*, n/a-n/a, doi:10.1002/asi.23323.

³⁴ La phrase a été prononcée, probablement pour la première fois, par Giovanni Battista Bachelet le 17 Juin 2008 dans le contexte d'une discussion parlementaire : http://documenti.camera.it/_dati/leg16/lavori/stencomm/07/audiz2/2008/0617/pdf002.pdf.

- Agasisti, T., & Catalano, G. (2006). Governance models of university systems—towards quasi-markets? Tendencies and perspectives: A European comparison. *Journal of Higher Education Policy and Management*, 28(3), 245-262.
- Amabile, F. (2012, 24/09). Suinicoltura nel cv dei ricercatori. *La Stampa*.
- Ammannati, L. (2013). Quis custodiet ipsos custodes? Qualche osservazione sul disegno della valutazione. In G. Colombini (Ed.), *Finanziamento, competizione ed accountability nel governo dell'università* (Vol. III, pp. 41-56). Napoli: ESI.
- ANVUR (2013). Rapporto finale. Valutazione della qualità della ricerca 2004-2010 (VQR 2004-2010). Roma: <http://www.anvur.org/rapporto/>.
- Aust, J., & Crespy, C. (2009). Napoléon renversé? Institutionnalisation des Pôles de recherche et d'enseignement supérieur et réforme du système académique français. *Revue française de science politique*, 59(5), 915-938.
- Baccini, A. (2010). *Valutare la ricerca scientifica. Uso e abuso degli strumenti bibliometrici*. Bologna: Il Mulino.
- Baccini, A. (2012). Chi vuoi sfrondare? Ecco l'isola delle riviste, il reality "scientifico" dell'ANVUR. Terzo episodio della trilogia delle riviste pazze. <http://www.roars.it/online/chi-vuoi-sfrondare-ecco-lisola-delle-riviste-il-reality-scientifico-dell'anvur-terzo-episodio-della-trilogia-delle-riviste-pazze/>.
- Baccini, A. (2013). Come e perché ridisegnare la valutazione. *Il Mulino*, 2013(1), 80-87, doi:10.1402/44138.
- Baccini, A. (2014). La VQR di Area 13: una riflessione di sintesi. *Statistica & Società*, 3(3), 32-37.
- Baccini, A., Coin, F., & Sirilli, G. (2013). Costi e benefici della valutazione della ricerca e della didattica. *Paradoxa*, VII(2), 46-61.
- Baccini, A., & De Nicolao, G. (2012). VQR: tutte le valutazioni sono uguali, ma alcune sono più uguali delle altre. <http://www.roars.it/online/vqr-tutte-le-valutazioni-sono-uguali-ma-alcune-sono-piu-uguali-delle-altre/>.
- Baccini, A., & De Nicolao, G. (2015). Do they agree? Bibliometric evaluation vs informed peer review in the Italian research assessment exercise. *arXiv:1505.00115*.
- Banfi, A. (2013). Prima che la nave affondi: un rapido bilancio della riforma dell'università e qualche possibile intervento correttivo. *Giornale di diritto amministrativo*, 5, 548-559.
- Banfi, A. (2015). Academic recruitment in Italy: an overview for foreigners. *Ius Publicum Network Review*, 1, 1-10.
- Banfi, A., & De Nicolao, G. (2012). Sesso, droga e chiesa: le pazze riviste ANVUR sempre più pazze (episodio 2 della trilogia). <http://www.roars.it/online/sesso-droga-e-chiesa-le-pazze-riviste-anvur-sempre-piu-pazze-episodio-2-della-trilogia/>.
- Banfi, A., & Viesti, G. (2015). Meriti e bisogni nel finanziamento del sistema universitario italiano. *Working papers RES* (pp. 63). Palermo: Fondazione RES.
- Barbati, C. (2013). L'autonomia universitaria "alla prova" della qualità e dell'efficienza. In G. Colombini (Ed.), *Finanziamento, competizione ed accountability nel governo dell'università* (Vol. 3, pp. 521-530). Napoli: ESI.
- Barbati, C., & Lenzi, A. (2011). Il Consiglio Universitario Nazionale: quale ruolo dopo la legge n. 240/2010? In M. Brollo, & R. De Luca Tamajo (Eds.), *La riforma dell'università tra legge e statuti. Analisi interdisciplinare della legge n. 240/2010* (pp. 319-322). Milano: Giuffrè.
- Benedetto, S. (2011). Atenei aperti al merito fin dal reclutamento. *IlSole24ore*.
- Bertocchi, G., Gambardella, A., Jappelli, T., Nappi, C. A., & Peracchi, F. (2015). Bibliometric evaluation vs. informed peer review: Evidence from Italy. *Research Policy*, 44(2), 451-466, doi:10.1016/j.respol.2014.08.004.
- Bezès, P. (2005). Le renouveau du contrôle des bureaucraties. L'impact du New Public Management. *Informations sociales*, 6(126), 26-37.
- Bleiklie, I., & Michelsen, S. (2013). Comparing HE policies in Europe. *Higher Education*, 65(1), 113-133, doi:10.1007/s10734-012-9584-6.
- Bonaccorsi, A. (2013). La valutazione di Bertoldo. *Il Mulino*(2), 257-266.
- Bonazzi, F. (2007, 22/01). I corrotti e i furbi ora li denuncio. Intervista a Fabio Mussi. *L'Espresso*.
- Calza, L., & Garbisa, S. (1995). Italian professorships. *Nature*, 374, 492.

- Capano, G. (2013). Il governo del sistema universitario italiano tra competizione regolata/coesione sistemica o frammentazione. In G. Capano, & M. Meloni (Eds.), *Il costo dell'ignoranza. L'università italiana e la sfida Euopra 2020* (pp. 71-93). Bologna: Il Mulino.
- Carlucci, D., & Castaldo, A. (2009). *Un paese di baroni*. Milano: Chiarelettere.
- Cassese, S. (2013). L'ANVUR ha ucciso la valutazione, viva la valutazione! *Il Mulino*(1), 73-79, doi:10.1402/44137.
- Checchi, D. (1999). Tenure. An appraisal of a national selection process for associate professorship. *Giornale degli Economisti e Annali di Economia*, 58(2), 137-181.
- Checchi, D. (2013). Valutazione: c'è un medicinale adatto? *Il Mulino*(2), 266-272.
- Coniglione, F. (2011). *Maledetta Università*. Trapani: Di Girolamo.
- De Nicolao, G. (2012). Mediane truccate? Analizziamo le prove. <http://www.roars.it/online/mediane-truccate-analizziamo-le-prove/>.
- De Nicolao, G. (2013). Abilitazioni: ecco le liste "last minute" per le riviste di classe A. <http://www.roars.it/online/abilitazioni-ecco-le-liste-last-minute-per-le-riviste-di-classe-a/comment-page-1/>.
- De Nicolao, G. (2014). Voti VQR ai dipartimenti: le normalizzazioni fai-da-te di CRUI e ANVUR. www.roars.it.
- De Nicolao, G., Banfi, A., & Baccini, A. (2013). Le abilitazioni a cucù! La nuova lista non c'è più! <http://www.roars.it/online/le-abilitazioni-a-cucu-la-nuova-lista-non-ce-piu/>.
- Della Cananea, G. (2011). Interpretazioni della costituzione e valutazioni della ricerca: interrogativi e ipotesi. *MUNUS. Rivista giuridica dei servizi pubblici*, 1(1), 633-642.
- Egghe, L. (2010). The Hirsch index and related impact measures. *Annual Review of Information Science and Technology*, 44(1), 65-114, doi:10.1002/aris.2010.1440440109.
- Fantoni, S. (2012, 20/12). Docenti e pubblicazioni universitarie. *Il Corriere della Sera*, p. 59.
- Ferrera, M. (2013, 25/02). Professori affetti da "benaltrismo" rallentano l'efficienza degli atenei. *Il Corriere della Sera*.
- Fiori, S. (2012, 04/02). Daremo le pagelle ai professori per fare la classifica delle università *La Repubblica*.
- Gagliarducci, S., Ichino, A., Peri, G., & Perotti, R. (2005). Lo splendido isolamento dell'università italiana. In T. Boeri, R. Faini, A. Ichino, G. Pisauro, & C. Scarpa (Eds.), *Oltre il Declino* (pp. 151-207). Bologna: Il Mulino.
- Garbisa, S., & Calzà, L. (1995). *Il peso della qualità accademica ovvero manuale per l'uso di parametri oggettivi*. Padova: CLUEP.
- Giavazzi, F. (2012). Se Bersani fa scuola. *lavoce.info*.
- Gingras, Y. (2008). Du mauvais usage de faux indicateurs. *Revue d'histoire moderne et contemporaine*, 5(4bis), 67-79.
- Graziosi, A. (2010). *L'università per tutti. Riforme e crisi del sistema universitario italiano*. Bologna: Il Mulino.
- Graziosi, A. (2012). Per una storia della terza mediana. <http://www.roars.it/online/per-una-storia-della-terza-mediana/>.
- Hess, D. J. (2012). *Science Studies: An Advanced Introduction* New York: New York University Press.
- Hicks, D. (2010). *Overview of models of performance-based research funding systems*. Paper presented at the Performance-based Funding for Public Research in Tertiary Education Institutions, Paris, 21 June 2010
- Hoggett, P. (1996). New modes of control in the public service. *Public Administration*, 74(1), 9-32, doi:10.1111/j.1467-9299.1996.tb00855.x.
- Lenzi, A., & Barbatì, C. (2011). Per un riordino del Consiglio Universitario Nazionale. *MUNUS. Rivista giuridica dei servizi pubblici*, 1(1), 675-680.
- Lozano, G. A., Larivière, V., & Gingras, Y. (2012). The weakening relationship between the impact factor and papers' citations in the digital age. *Journal of the American Society for Information Science and Technology*, 63(11), 2140-2145, doi:10.1002/asi.22731.
- Massarenti, A. (2012, 30/09). Ma il principio è da salvare. *IlSole24ore*.
- Mazzotti, M. (2012, 08/11). Listing wildly. *Times Higher Education*.
- McNay, I. (2007). Research assessment; research autonomy. In C. Kayrooz, G. S. Akerlind, & M. Tight (Eds.), *Autonomy in Social Science Research. International Perspectives on Higher Education Research* (Vol. 4, pp. 183-216). Oxford: Emerald.

- McNay, I. (2011). Research assessment: work in progress, or 'la lotta continua' In M. Saunders, P. Trowler, & V. Bamber (Eds.), *Reconceptualising Evaluation in Higher Education The Practice Turn* (pp. 51-57). New York: McGraw Hill.
- Modica, L. (2012). Passato e futuro della valutazione della ricerca universitaria. In P. Miccoli, & A. Fabris (Eds.), *Valutare la ricerca. Capire, applicare, difendersi*. (pp. 13-34). Pisa: ETS publisher.
- Moretti, M. (2009). Un pamphlet truccato. *Allegoria*, XXI(59), 201-214.
- Moretti, M. (2010). Sul governo dell'università nell'Italia contemporanea. *Annali di Storia delle Università Italiane*, 14.
- OECD (2002). *Frascati Manual 2002. Proposed Standard Practice for Surveys on Research and Experimental Design*. Paris: OECD Publishing.
- Perotti, R. (2008). *L'università truccata*. Torino: Einaudi.
- Pezzella, A. (2011). Il Consiglio Universitario Nazionale dopo la riforma Gelmini. *menodizero*, 2(6).
- Pinelli, C. (2011). L'autonomia universitaria e la valutazione dell'attività scientifica. *MUNUS. Rivista giuridica dei servizi pubblici*, 1(1), 567-588.
- Pinto, V. (2013). La valutazione come strumento di intelligence e tecnologia di governo. *Aut Aut*(360), 16-42.
- Pontille, D., & Torny, D. (2010). The controversial policies of journal ratings: Evaluating social sciences and humanities. *Research Evaluation*, 19(5), 347-360.
- Reale, E. (Ed.). (2008). *La valutazione della ricerca pubblica. Una analisi della valutazione triennale della ricerca*. Milano: Angeli.
- Reale, E. (2012). *Shifting power in higher education policy making. The influence of the evaluation agencies for transforming universities knowledge production*. Paper presented at the 28th EGOS Colloquium, Helsinki,
- Rebora, G. (2013). *Nessuno mi può giudicare? L'università e la valutazione*. Milano: Guerini e Associati.
- Redazione Roars (2012a). Le riviste "scientifiche" dell'ANVUR: dal sacro al profano e dalle stelle alle stalle. <http://www.roars.it/online/le-riviste-scientifiche-dellanvur-dal-sacro-al-profano-e-dalle-stelle-alle-stalle/>.
- Redazione Roars (2012b). Mediane ANVUR: le dieci domande di ROARS. <http://www.roars.it/online/mediane-anvur-le-dieci-domande-di-roars/>.
- Redazione Roars (2015). *Università 3.0. Quattro anni vissuti pericolosamente*. Roma: ecommons.
- Regini, M. (2009). *Malata e denigrata. L'Università italiana a confronto con l'Europa*. Roma: Donzelli.
- Rocca, G. (2014). *Riprendiamo in mano il nostro destino. Relazione del Presidente*. Paper presented at the Assemblea generale di Assolombarda, Milano, 9/6/2014
- Rodrik, D. (2008). Second-Best Institutions. *American Economic Review*, 98(2), 100-104, doi:10.1257/aer.98.2.100.
- Rubele, R. (2012). Appunti per una storia dell'ANVUR (I-III). In roars (Ed.), www.roars.it.
- Sandulli, A. (2011). La rappresentanza dei saperi e delle discipline: il Consiglio Universitario Nazionale e la teoria del doppio centro. *MUNUS. Rivista giuridica dei servizi pubblici*, 1(3), 599-622.
- Seglen, P. O. (1997). Why the impact factor for journals should not be used for evaluating research. *BMJ*, 314, 497.
- Stella, G. A. (2012, 17/10). Vuoi diventare professore? Scrivi su Suinicoltura. *Il Corriere della Sera*.
- Sylos Labini, F., & Zapperi, S. (2010). *I ricercatori non crescono sugli alberi*. Bari: Laterza.
- Travi, A. (2011). Organi consultivi e processi decisionali dopo la riforma universitaria. *MUNUS. Rivista giuridica dei servizi pubblici*, 1(1), 623-632.
- Vaira, M. (2011). Evaluation as ceremony in the Italian university system. In M. Saunders, P. Trowler, & V. Bamber (Eds.), *Reconceptualising Evaluation in Higher Education The Practice Turn* (pp. 40-45). New York: McGraw Hill.
- Zunino, C. (2012, 27/09). Il gotha delle riviste scientifiche oggi. *La Repubblica*.

Tableau 1. Produits évalués par domaine scientifique.

GEV	Discipline	Nombre des produits des recherches attendues	Nombre des produits soumis à évaluation	% <i>manquant</i>
01	Mathématique et informatique	11 752	10 685	9,1
02	Physique	20 286	19 773	2,5
03	Chimie	11 933	11 608	2,7
04	Sciences de la terre	8 859	8 433	4,8
05	Biologie	17 268	16 407	5,0
06	Médecine	29 454	26 713	9,3
07	Sciences agraires et vétérinaires	10 349	10 004	3,3
08	Ingénierie civile et architecture	9 934	9 533	4,0
09	Ingénierie industrielle et de l'information	16 858	16 347	3,0
10	Sc. de l'antiquité, philologie, littérature e histoire de l'art	14 637	14 073	3,9
11	Histoire, philosophie, psychologie, pédagogie	13 487	13 152	2,5
12	Droit	12 798	11 882	7,2
13	Economie et statistique	12 654	11 941	5,6
14	Science politique et sociologie	4 494	4 327	3,7
Totale		194.763	184.878	5,1

Source : ANVUR

Tableau 2. La variété des méthodes d'évaluation adoptées par l'ANVUR pour la VQR

Denomination du Group d'Experts de evaluation (GEV)	Discipline	Class.	Nombre de sous-groupes disciplinaires	Database	Nombre de listes des revues	Methode de categorization des revues	Nombre des matrices utilisés
1	Mathématique et informatique	B	4	WoS; Scopus; MathScinet	6	propriétaire	6
2	Physique	B	4	WoS; Scopus		IF ajusté	4
3	Chimie	B	3	WoS	3	IF	6
4	Sciences de la terre	B	1	WoS; Scopus		IF ajusté	2
5	Biologie	B	4	WoS		IF	8
6	Médecine	B	4	WoS		IF	8
7	Sciences agraires et vétérinaires	B	2	WoS; Scopus		IF ajusté	2
8	Ingénierie civile et architecture	B/NB	2	WoS; Scopus	1	propriétaire	4
9	Ingénierie industrielle et de l'information	B	3	WoS; Scopus		ACP	8
10	Sc. de l'antiquité, philologie, littérature e histoire de l'art	NB	5	-		Non Biblio	-
11	Histoire, philosophie, psychologie, pédagogie	NB/B	4	WoS		Non Biblio/IF ajusté	-
12	Droit	NB	3	-		Non Biblio	-
13	Economie et statistique	NB	3	WoS; Scopus; Google Scholar	1	Biblio	-
14	Science politique et sociologie	NB	2	WoS; Scopus; Google Scholar		Non Biblio	-
Total			44				48

Légende. *Class.*: «B» discipline bibliométrique; « NB » discipline non-bibliométrique; *Database*. « WoS » : Web of Science, Thomson Reuters, <http://thomsonreuters.com/web-of-science/>; « Scopus »: Scopus database, Elsevier, <http://www.scopus.com>; « MathSciNet »: MathSciNet database, American Mathematical Society, <http://www.ams.org/mathscinet/>; «Google Scholar», Google Inc. <http://scholar.google.com/>. *Méthode de classification des revues*. «IF ajusté»: Impact Factor selon WoS avec ajustements développés par le GEV ; «propriétaire» : ranking développé directement par le GEV ; «ACP» : analyse en composantes principales ; «Non Biblio» : revues catégorisées par un jury nommé par ANVUR ; «Biblio »: catégorisation des revues sur données bibliométriques.

Source. Elaboration sur données ANVUR.